

RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical
Information Center (STIC) no errors detected.

Application Serial Number: 10/076,157A
Source: 1FW/6
Date Processed by STIC: 10/22/04

ENTERED



IFW16

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/076,157A

DATE: 10/22/2004

TIME: 12:20:07

Input Set : A:\076157.txt

Output Set: N:\CRF4\10222004\J076157A.raw

```

3 <110> APPLICANT: Pompejus, Markus
4     Seulberger, Harald
5     Hoeffken, Hans Wolfgang
6     Doval, Jose Luis Revuelta
7     Jimenez, Alberto
8     Garcia, Maria Angeles Santos
10 <120> TITLE OF INVENTION: Phosphoriboxyl-Pyrophosphate Synthetase Polypeptide
12 <130> FILE REFERENCE: PF48687-2/DP
14 <140> CURRENT APPLICATION NUMBER: US 10/076,157A
15 <141> CURRENT FILING DATE: 2002-02-15
17 <150> PRIOR APPLICATION NUMBER: Germany, 19757755.5
18 <151> PRIOR FILING DATE: 1997-12-23
20 <160> NUMBER OF SEQ ID NOS: 21
22 <170> SOFTWARE: WordPerfect 8
24 <210> SEQ ID NO: 1
25 <211> LENGTH: 1911
26 <212> TYPE: DNA
27 <213> ORGANISM: Ashbya gossypii
29 <400> SEQUENCE: 1
C--> 31 ggtagtcgct catcgacaga cacaatcgcg tgttctctct gaatcggtcca ttgggtgtca      60
33 gcacctgat cgcgggcgga tgggaatgggt aatcattagg aaacaccaat gtcccatggg      120
35 attgtccgct ctcgtatggt gtctcaggag gaccctgat cacgtagtgc cacaccagga      180
37 tattgtcttc ctttgggtgt gccacgatgt agggcggggg gttctcggtc atcattttgt      240
39 actcctttga gagccgcttg tacgcctgtc ttgatgccat cttgcctact attagtttct      300
41 caccacttcc cgccaaacaa tctgcacttt acgagcgcta tctatccctc gggtcgctct      360
43 agttgattat tggcgaaaact gatagttcag gtacttccat gatgcggtca tatccacgta      420
45 tgtgatcacg tgatcatcag ccatgctgcc agctcacggg cctgcctaca ctattggagg      480
47 ctctgtgagt catgatattat tgcatatcaa gccagatag tcgttgggga tactaccgtt      540
49 gccgcgatga gctccgatat taagttgtag ccaaaaattt taacggatga cttcttaaca      600
51 gttattgacg ccgcaatcct acgcc atg tcg tcc aat agc ata aag ctg cta      652
52                                     Met Ser Ser Asn Ser Ile Lys Leu Leu
53                                     1                               5
55 gca ggt aac tcg cac ccg gac cta gct gag aag gtc tcc gtt cgc cta      700
56 Ala Gly Asn Ser His Pro Asp Leu Ala Glu Lys Val Ser Val Arg Leu
57 10                               15                               20                               25
59 ggt gta cca ctt tcg aag att gga gtg tat cac tac tct aac aaa gag      748
60 Gly Val Pro Leu Ser Lys Ile Gly Val Tyr His Tyr Ser Asn Lys Glu
61                               30                               35                               40
63 acg tca gtt act atc ggc gaa agt atc cgt gat gaa gat gtc tac atc      796
64 Thr Ser Val Thr Ile Gly Glu Ser Ile Arg Asp Glu Asp Val Tyr Ile
65                               45                               50                               55
67 atc cag aca gga acg ggg gag cag gaa atc aac gac ttc ctc atg gaa      844
68 Ile Gln Thr Gly Thr Gly Glu Gln Glu Ile Asn Asp Phe Leu Met Glu

```

RAW SEQUENCE LISTING

DATE: 10/22/2004

PATENT APPLICATION: US/10/076,157A

TIME: 12:20:07

Input Set : A:\076157.txt

Output Set: N:\CRF4\10222004\J076157A.raw

69	60	65	70	
71	ctg ctc atc atg atc cat gcc tgc cgg tca gcc tct gcg cgg aag atc	892		
72	Leu Leu Ile Met Ile His Ala Cys Arg Ser Ala Ser Ala Arg Lys Ile			
73	75 80 85			
75	aca gcg gtt ata cca aac ttc cct tac gca aga caa gac aaa aag gac	940		
76	Thr Ala Val Ile Pro Asn Phe Pro Tyr Ala Arg Gln Asp Lys Lys Asp			
77	90 95 100 105			
79	aag tgc cga gca ccg ata act gcc aag ctg gtg gcc aag atg cta gag	988		
80	Lys Ser Arg Ala Pro Ile Thr Ala Lys Leu Val Ala Lys Met Leu Glu			
81	110 115 120			
83	acc gcg ggg tgc aac cac gtt atc acg atg gat ttg cac gcg tct caa	1036		
84	Thr Ala Gly Cys Asn His Val Ile Thr Met Asp Leu His Ala Ser Gln			
85	125 130 135			
87	att cag ggt ttc ttc cac att cca gtg gac aac cta tat gca gag ccg	1084		
88	Ile Gln Gly Phe Phe His Ile Pro Val Asp Asn Leu Tyr Ala Glu Pro			
89	140 145 150			
91	aac atc ctg cac tac atc caa cat aat gtg gac ttc cag aat agt atg	1132		
92	Asn Ile Leu His Tyr Ile Gln His Asn Val Asp Phe Gln Asn Ser Met			
93	155 160 165			
95	ttg gtc gcg cca gac gcg ggg tgc gcg aag cgc acg tgc acg ctt tgc	1180		
96	Leu Val Ala Pro Asp Ala Gly Ser Ala Lys Arg Thr Ser Thr Leu Ser			
97	170 175 180 185			
99	gac aag ctg aat ctc aac ttc gcg ttg atc cac aaa gaa cgg cag aag	1228		
100	Asp Lys Leu Asn Leu Asn Phe Ala Leu Ile His Lys Glu Arg Gln Lys			
101	190 195 200			
103	gcg aac gag gtc tgc cgg atg gtg ttg gtg ggt gat gtc gcc gac aag	1276		
104	Ala Asn Glu Val Ser Arg Met Val Leu Val Gly Asp Val Ala Asp Lys			
105	205 210 215			
107	tcc tgt att att gta gac gac atg gcg gac acg tgc gga acg cta gtg	1324		
108	Ser Cys Ile Ile Val Asp Asp Met Ala Asp Thr Cys Gly Thr Leu Val			
109	220 225 230			
111	aag gcc act gac acg ctg atc gaa aat tgt gcg aaa gaa gtg att gcc	1372		
112	Lys Ala Thr Asp Thr Leu Ile Glu Asn Cys Ala Lys Glu Val Ile Ala			
113	235 240 245			
115	att gtg aca cac ggt ata ttt tct ggc ggc gcc cgc gag aag ttg cgc	1420		
116	Ile Val Thr His Gly Ile Phe Ser Gly Gly Ala Arg Glu Lys Leu Arg			
117	250 255 260 265			
119	aac agc aag ctg gca cgg atc gta agc aca aat acg gtg cca gtg gac	1468		
120	Asn Ser Lys Leu Ala Arg Ile Val Ser Thr Asn Thr Val Pro Val Asp			
121	270 275 280			
123	ctc aat cta gat atc tac cac caa att gac att agt gcc att ttg gcc	1516		
124	Leu Asn Leu Asp Ile Tyr His Gln Ile Asp Ile Ser Ala Ile Leu Ala			
125	285 290 295			
127	gag gca att aga agg ctt cac aac ggg gaa agt gtg tgc tac ctg ttc	1564		
128	Glu Ala Ile Arg Arg Leu His Asn Gly Glu Ser Val Ser Tyr Leu Phe			
129	300 305 310			
131	aat aac gct gtc atg tagtgctgtc agtggcagat gcatgatcgc tggcctaatt	1619		
132	Asn Asn Ala Val Met			
133	315			

RAW SEQUENCE LISTING

DATE: 10/22/2004

PATENT APPLICATION: US/10/076,157A

TIME: 12:20:07

Input Set : A:\076157.txt

Output Set: N:\CRF4\10222004\J076157A.raw

```

135 atctgtgtaa gttgatacaa tgcagtaa acagtacata aaactgaatg tttttcactt 1679
137 aggggtgctt tgttgttctg atagcgtgtg tgcgaatttg gaggtgaaag ttgaacatca 1739
139 cgtaatgaat acaaacaaga ttgcacatta ggaaaagcga taaattatattt attatttgca 1799
141 actggccttt gagcgtttaa gcctgaacat ttttgccctt ttgtttgacc gtaccgttat 1859
143 cactcgctct tatatatggc tatecttctc ttccggaact tcttcgagcg ta 1911
146 <210> SEQ ID NO: 2
147 <211> LENGTH: 318
148 <212> TYPE: PRT
149 <213> ORGANISM: Ashbya gossypii
151 <400> SEQUENCE: 2
153 Met Ser Ser Asn Ser Ile Lys Leu Leu Ala Gly Asn Ser His Pro Asp
154 1 5 10 15
156 Leu Ala Glu Lys Val Ser Val Arg Leu Gly Val Pro Leu Ser Lys Ile
157 20 25 30
159 Gly Val Tyr His Tyr Ser Asn Lys Glu Thr Ser Val Thr Ile Gly Glu
160 35 40 45
162 Ser Ile Arg Asp Glu Asp Val Tyr Ile Ile Gln Thr Gly Thr Gly Glu
163 50 55 60
165 Gln Glu Ile Asn Asp Phe Leu Met Glu Leu Leu Ile Met Ile His Ala
166 65 70 75 80
168 Cys Arg Ser Ala Ser Ala Arg Lys Ile Thr Ala Val Ile Pro Asn Phe
169 85 90 95
171 Pro Tyr Ala Arg Gln Asp Lys Lys Asp Lys Ser Arg Ala Pro Ile Thr
172 100 105 110
174 Ala Lys Leu Val Ala Lys Met Leu Glu Thr Ala Gly Cys Asn His Val
175 115 120 125
177 Ile Thr Met Asp Leu His Ala Ser Gln Ile Gln Gly Phe Phe His Ile
178 130 135 140
180 Pro Val Asp Asn Leu Tyr Ala Glu Pro Asn Ile Leu His Tyr Ile Gln
181 145 150 155 160
183 His Asn Val Asp Phe Gln Asn Ser Met Leu Val Ala Pro Asp Ala Gly
184 165 170 175
186 Ser Ala Lys Arg Thr Ser Thr Leu Ser Asp Lys Leu Asn Leu Asn Phe
187 180 185 190
189 Ala Leu Ile His Lys Glu Arg Gln Lys Ala Asn Glu Val Ser Arg Met
190 195 200 205
192 Val Leu Val Gly Asp Val Ala Asp Lys Ser Cys Ile Ile Val Asp Asp
193 210 215 220
195 Met Ala Asp Thr Cys Gly Thr Leu Val Lys Ala Thr Asp Thr Leu Ile
196 225 230 235 240
198 Glu Asn Cys Ala Lys Glu Val Ile Ala Ile Val Thr His Gly Ile Phe
199 245 250 255
201 Ser Gly Gly Ala Arg Glu Lys Leu Arg Asn Ser Lys Leu Ala Arg Ile
202 260 265 270
204 Val Ser Thr Asn Thr Val Pro Val Asp Leu Asn Leu Asp Ile Tyr His
205 275 280 285
207 Gln Ile Asp Ile Ser Ala Ile Leu Ala Glu Ala Ile Arg Arg Leu His
208 290 295 300
210 Asn Gly Glu Ser Val Ser Tyr Leu Phe Asn Asn Ala Val Met

```

RAW SEQUENCE LISTING

DATE: 10/22/2004

PATENT APPLICATION: US/10/076,157A

TIME: 12:20:07

Input Set : A:\076157.txt

Output Set: N:\CRF4\10222004\J076157A.raw

```

211 305          310          315
214 <210> SEQ ID NO: 3
215 <211> LENGTH: 5369
216 <212> TYPE: DNA
217 <213> ORGANISM: Ashbya gossypii
219 <400> SEQUENCE: 3
C--> 222 aagcttgacc ttggctggca cttgagtcgg cagacagggtg gactaacccg agca atg      57
223                                     Met
224                                     1
226 gat cgt ggt tgt aaa ggt atc tct tat gtg ctc agt gca atg gtt ttt      105
227 Asp Arg Gly Cys Lys Gly Ile Ser Tyr Val Leu Ser Ala Met Val Phe
228          5          10          15
230 cac ata ata ccg att aca ttt gaa ata tcg atg gta tgt ggc ata ttg      153
231 His Ile Ile Pro Ile Thr Phe Glu Ile Ser Met Val Cys Gly Ile Leu
232          20          25          30
234 aca tac cag ttt ggt gct tcc ttc gct gct ata aca ttc tcg act atg      201
235 Thr Tyr Gln Phe Gly Ala Ser Phe Ala Ala Ile Thr Phe Ser Thr Met
236          35          40          45
238 ctt ctt tac tcc atc ttt act ttc aga acg acg gcg tgg cgc aca cgg      249
239 Leu Leu Tyr Ser Ile Phe Thr Phe Arg Thr Thr Ala Trp Arg Thr Arg
240 50          55          60          65
242 ttt agg cgt gat gcg aac aag gct gac aat aag gcc gct agt gtg gca      297
243 Phe Arg Arg Asp Ala Asn Lys Ala Asp Asn Lys Ala Ala Ser Val Ala
244          70          75          80
246 ttg gat tcc cta ata aat ttt gaa gct gta aag tat ttc aat aac gag      345
247 Leu Asp Ser Leu Ile Asn Phe Glu Ala Val Lys Tyr Phe Asn Asn Glu
248          85          90          95
250 aag tac ctt gcg gac aag tat cac aca tcc ttg atg aag tac cgg gat      393
251 Lys Tyr Leu Ala Asp Lys Tyr His Thr Ser Leu Met Lys Tyr Arg Asp
252          100          105          110
254 tcc cag ata aag gtc tcg caa tcg ctg gcg ttt ttg aac acc ggc cag      441
255 Ser Gln Ile Lys Val Ser Gln Ser Leu Ala Phe Leu Asn Thr Gly Gln
256          115          120          125
258 aac cta att ttt acc act gca ctg act gca atg atg tat atg gcc tgt      489
259 Asn Leu Ile Phe Thr Thr Ala Leu Thr Ala Met Met Tyr Met Ala Cys
260 130          135          140          145
262 aat ggt gtt atg cag ggc tct ctt aca gtg ggg gat ctt gtg tta att      537
263 Asn Gly Val Met Gln Gly Ser Leu Thr Val Gly Asp Leu Val Leu Ile
264          150          155          160
266 aat caa ctg gta ttc cag ctc tcc gtg cca cta aac ttc ctt ggt agc      585
267 Asn Gln Leu Val Phe Gln Leu Ser Val Pro Leu Asn Phe Leu Gly Ser
268          165          170          175
270 gtc tac cgt gat ctc aag cag tct ctg ata gat atg gaa tct tta ttt      633
271 Val Tyr Arg Asp Leu Lys Gln Ser Leu Ile Asp Met Glu Ser Leu Phe
272          180          185          190
274 aaa ctg caa aaa aat cag gtc aca att aag aac tcc cca aat gcc cag      681
275 Lys Leu Gln Lys Asn Gln Val Thr Ile Lys Asn Ser Pro Asn Ala Gln
276          195          200          205
278 aac cta cca ata cac aaa ccg ttg gat att cgc ttt gaa aat gtt acg      729

```

RAW SEQUENCE LISTING

DATE: 10/22/2004

PATENT APPLICATION: US/10/076,157A

TIME: 12:20:07

Input Set : A:\076157.txt

Output Set: N:\CRF4\10222004\J076157A.raw

```

279 Asn Leu Pro Ile His Lys Pro Leu Asp Ile Arg Phe Glu Asn Val Thr
280 210          215          220          225
282 ttt ggc tat gac ccg gag cgg cgt ata ttg aac aat gtt tcg ttt acc      777
283 Phe Gly Tyr Asp Pro Glu Arg Arg Ile Leu Asn Asn Val Ser Phe Thr
284          230          235          240
286 atc cca gct gga atg aag act gcc ata gta ggc cca tcg ggc tcg ggg      825
287 Ile Pro Ala Gly Met Lys Thr Ala Ile Val Gly Pro Ser Gly Ser Gly
288          245          250          255
290 aag tcc acc att ttg aag ctc gta ttt aga ttc tat gag ccc gag caa      873
291 Lys Ser Thr Ile Leu Lys Leu Val Phe Arg Phe Tyr Glu Pro Glu Gln
292          260          265          270
294 ggt cgt atc cta gtt ggc ggc aca gat atc cgc gat tta gac ttg ctt      921
295 Gly Arg Ile Leu Val Gly Gly Thr Asp Ile Arg Asp Leu Asp Leu Leu
296          275          280          285
298 tct tta cgg aag gct atc ggt gtc gtg ccc caa gat act cct ctc ttc      969
299 Ser Leu Arg Lys Ala Ile Gly Val Val Pro Gln Asp Thr Pro Leu Phe
300 290          295          300          305
302 aat gac aca atc tgg gag aat gtt aaa ttc ggc aat atc agt tcc tct      1017
303 Asn Asp Thr Ile Trp Glu Asn Val Lys Phe Gly Asn Ile Ser Ser Ser
304          310          315          320
306 gac gat gag att ctc agg gcc ata gaa aaa gct caa ctc acg aag cta      1065
307 Asp Asp Glu Ile Leu Arg Ala Ile Glu Lys Ala Gln Leu Thr Lys Leu
308          325          330          335
310 ctc cag aac cta cca aag ggc gct tcc acc gtt gta ggg gag cgc ggt      1113
311 Leu Gln Asn Leu Pro Lys Gly Ala Ser Thr Val Val Gly Glu Arg Gly
312          340          345          350
314 ttg atg atc agc gga ggt gag aaa caa agg ctt gct att gct cgt gtg      1161
315 Leu Met Ile Ser Gly Gly Glu Lys Gln Arg Leu Ala Ile Ala Arg Val
316          355          360          365
318 ctt ttg aag gac gct ccg ctg atg ttt ttc gac gag gct aca agt gct      1209
319 Leu Leu Lys Asp Ala Pro Leu Met Phe Phe Asp Glu Ala Thr Ser Ala
320 370          375          380          385
322 ctg gat aca cac aca gag cag gca ctc ttg cac acc att cag cag aac      1257
323 Leu Asp Thr His Thr Glu Gln Ala Leu Leu His Thr Ile Gln Gln Asn
324          390          395          400
326 ttt tct tcc aat tca aag acg agc gtt tac gtt gcc cat aga ctg cgc      1305
327 Phe Ser Ser Asn Ser Lys Thr Ser Val Tyr Val Ala His Arg Leu Arg
328          405          410          415
330 aca atc gct gat gca gat aag atc att gtt ctt gaa caa ggt tct gtc      1353
331 Thr Ile Ala Asp Ala Asp Lys Ile Val Leu Glu Gln Gly Ser Val
332          420          425          430
334 cgc gaa gag ggc aca cac agc tcg ctg tta gcg tca caa gga tcc cta      1401
335 Arg Glu Glu Gly Thr His Ser Ser Leu Leu Ala Ser Gln Gly Ser Leu
336          435          440          445
338 tac cgg ggt ctg tgg gat att cag gaa aac cta acg ctt ccg gaa cgg      1449
339 Tyr Arg Gly Leu Trp Asp Ile Gln Glu Asn Leu Thr Leu Pro Glu Arg
340 450          455          460          465
342 cct gag cag tca acc gga tct cag cat gca tagacgtctg actagagatt      1499
343 Pro Glu Gln Ser Thr Gly Ser Gln His Ala

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/076,157A

DATE: 10/22/2004

TIME: 12:20:08

Input Set : A:\076157.txt

Output Set: N:\CRF4\10222004\J076157A.raw

L:31 M:112 C: (48) String data converted to lower case,
M:112 Repeated in SeqNo=1
L:222 M:112 C: (48) String data converted to lower case,
M:112 Repeated in SeqNo=3
L:356 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:496 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:500 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:504 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:508 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:512 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:516 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:520 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:524 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:528 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:532 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:536 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:540 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:544 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:548 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:552 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:556 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:560 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:564 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:568 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:572 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:576 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:580 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:584 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:588 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:902 M:112 C: (48) String data converted to lower case,
M:112 Repeated in SeqNo=7
L:976 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:7
L:1229 M:112 C: (48) String data converted to lower case,
M:112 Repeated in SeqNo=10
L:1515 M:112 C: (48) String data converted to lower case,
M:112 Repeated in SeqNo=12